

UC Berkeley

UC Berkeley Previously Published Works

Title

Long-oligomer microarray profiling in *Neurospora crassa* reveals the transcriptional program underlying biochemical and physiological events of conidial germination.

Permalink

<https://escholarship.org/uc/item/1c46351f>

Journal

Nucleic acids research, 33(20)

ISSN

0305-1048

Authors

Kasuga, Takao
Townsend, Jeffrey P
Tian, Chaoguang
et al.

Publication Date

2005

DOI

10.1093/nar/gki953

Peer reviewed

Long-oligomer microarray profiling in *Neurospora crassa* reveals the transcriptional program underlying biochemical and physiological events of conidial germination

Takao Kasuga¹, Jeffrey P. Townsend^{1,2}, Chaoguang Tian¹, Luz B. Gilbert¹, Gertrud Mannhaupt³, John W. Taylor¹ and N. Louise Glass^{1,*}

¹Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720-3102, USA, ²Department of Molecular and Cell Biology, University of Connecticut, Storrs, CT 06269, USA and ³Institute for Bioinformatics (MIPS), GSF National Research Center for Environment and Health, D-85764 Neuherberg, Germany

Received June 24, 2005; Revised September 22, 2005; Accepted October 19, 2005

ABSTRACT

To test the inferences of spotted microarray technology against a biochemically well-studied process, we performed transcriptional profiling of conidial germination in the filamentous fungus, *Neurospora crassa*. We first constructed a 70 base oligomer microarray that assays 3366 predicted genes. To estimate the relative gene expression levels and changes in gene expression during conidial germination, we analyzed a circuit design of competitive hybridizations throughout a time course using a Bayesian analysis of gene expression level. Remarkable consistency of mRNA profiles with previously published northern data was observed. Genes were hierarchically clustered into groups with respect to their expression profiles over the time course of conidial germination. A functional classification database was employed to characterize the global picture of gene expression. Consensus motif searches identified a putative regulatory component associated with genes involved in ribosomal biogenesis. Our transcriptional profiling data correlate well with biochemical and physiological processes associated with conidial germination and will facilitate functional predictions of novel genes in *N.crassa* and other filamentous ascomycete species. Furthermore, our dataset on conidial germination allowed comparisons to transcriptional mechanisms associated with germination processes of diverse propagules, such as teliospores of the phytopathogenic fungus

Ustilago maydis and spores of the social amoeba *Dictyostelium discoideum*.

INTRODUCTION

In the past 10 years, ~75 eukaryotic genomes have been completed or assembled, yet molecular functions of a large number of genes are still unknown [e.g. 13% of genes in *Homo sapiens* (1) and 38% in *Saccharomyces cerevisiae* (Saccharomyces Genome Database, <http://www.yeastgenome.org/>). Gene disruption, proteomics and mRNA transcription profiling have been extensively used in *S.cerevisiae* to characterize gene functions. The simple life cycle and small genome size (12 Mb) make *S.cerevisiae* an experimentally powerful model eukaryotic organism. However, *S.cerevisiae* provides limited insight into the complex developmental programs and inter-cellular communication observed in multicellular eukaryotic species. *Neurospora crassa*, a multicellular filamentous ascomycete fungus, has been used as a model laboratory organism for over 70 years and is a promising model system for addressing questions about multicellular development.

An abundance of tools and mutants for genetic, molecular genetics and biochemical analyses of *N.crassa* are available (2,3). *N.crassa* is characterized by a complex lifecycle; there are at least 28 distinct cell types (4,5). The DNA sequence of the 40 Mb genome of *N.crassa* encodes ~10 000 genes (6). Functional annotation of these genes has succeeded largely on the basis of computer predictions using patterns recognized in characterized genes and homology searches to proteins in public protein databases (4,6,7). However, over 40% of the *N.crassa* genes have no counterpart in public databases and are listed as 'hypothetical'.

*To whom correspondence should be addressed. Tel: +1 510 643 2399; Fax: +1 510 642 4995; Email: Lglass@nature.berkeley.edu

The advent of microarray technology makes it possible to measure transcription in thousands of genes simultaneously by hybridizing labeled transcripts (targets) to probes attached to the surface of microscope slides (8,9). Transcriptional profiling data have been used to predict biological function of novel genes in *S.cerevisiae* (10). Gene expression data generated by microarray technology may be used to expedite the annotation process of the *N.crassa* genome and will aid in assigning putative functions for unique genes. However, most physiological and cellular phenomena are driven directly by proteins, rather than by genes or mRNAs. The total amount and/or activity of proteins in the cell is controlled through pre-, co- and post-translational modification, in addition to mRNA abundance. Accordingly, it has been argued that mRNA abundance, discernable by microarrays, is a poor indicator of protein abundance or protein activity [reviewed in (11,12)]. If mRNA abundance had little effect upon protein or cellular activity, then mRNA profiling would provide little insight into underlying cellular mechanisms. To test this hypothesis of non-concordance, we chose to evaluate gene expression patterns associated with asexual spore (conidial) germination of *N.crassa*. This developmental process is ideal for testing this hypothesis, because morphological and biochemical aspects associated with conidial germination have been studied in great detail (13,14) and because conidial germination is a morphologically simple, rapid and a highly important developmental process in this fungus. If the hypothesis were to be correct, discordance between mRNA profiles and morphological/biochemical observations would be revealed.

N.crassa, similar to numerous other filamentous fungi, forms a mycelial colony via hyphal tip growth, branching and hyphal fusion. Upon exhaustion of nutrients, *N.crassa* produces numerous asexual multinucleate spores, called macroconidia, which bud from the tips of specialized aerial hyphae, or conidiophores (2,15). In addition to macroconidia, *N.crassa* also produces uninucleate microconidia, which are extruded directly from cells of microconidiophores (16) and arthroconidia, which are derived by segmentation (2). In this report, we use 'conidia' to refer exclusively to macroconidia. In other filamentous ascomycete species, conidia are often the infective propagules associated with both plant and animal mycoses, e.g. rice blast, caused by *Magnaporthe grisea*, and aspergilliosis, caused by *Aspergillus fumigatus*, which is a common and frequently lethal infection of individuals with a compromised immune system. It has been proposed that the identification and targeting of conidial germination-specific processes provides an excellent strategy for drug and fungicide development in these pathogenic fungi (17).

Conidial germination in filamentous fungi is a highly regulated process that is triggered by environmental stimuli. Dehydrated conidia remain viable for many years, a resilience which is lost soon after germination begins (2). When *N.crassa* conidia are put in water with salts and a carbon source (13), the outer, hydrophobic rodlet layer of the cell wall is lost and the conidia swell (18). In the first 10 min of exposure to liquid medium, free glutamic acid is degraded; trehalose is mobilized; and aspartic acid and γ -aminobutyric acid are formed, followed by the initiation of protein and RNA synthesis (13,19–22). The first morphological sign of conidial germination is the formation of the germ tube, which requires polarity establishment (23). Coincident with germ tube formation,

DNA synthesis and nuclear division is initiated and mitochondrial oxidative phosphorylation can be detected. High levels of arginine and ornithine accumulate and synthesis of chitin (a component of the fungal cell wall) occurs. In *N.crassa*, the germ tube elongates $\sim 200\ \mu\text{m}$ before the initiation of branching. Germ tubes or conidia in close proximity often undergo fusion via specialized hyphae, called conidial anastomosis tubes (14,24). In Vogel's medium (25), maximal growth rate is established soon after branching is initiated, $\sim 8\ \text{h}$ post-inoculation.

Although biochemical aspects associated with conidial germination have been well documented in *N.crassa*, fundamental genetic mechanisms, such as those that drive the germination process and underlie the timing of gene expression and metabolic pathway activation, remain obscure. In this study, we tested two hypotheses that relate gene expression levels to their fundamental biological impact. The first hypothesis tested directly whether established timing of biochemical events and measurements of enzyme activity during conidial germination were consistent with our transcriptional profiling dataset. These analyses identified expression profiles of a large number of genes which lacked any previous functional annotation, proving the utility of microarray data for annotation of the *Neurospora* genome. The second hypothesis tested whether transcriptional profiles associated with conidial germination shared similarities to transcriptional profiles obtained from spore germination from the phylogenetically distant species, such as *Ustilago maydis*, a basidiomycete fungus and *Dictyostelium discoideum*, an amoeboid protist species.

MATERIALS AND METHODS

Construction of *N.crassa* oligonucleotide microarrays

We chose to construct oligomer microarrays for transcriptional profiling experiments for *N.crassa*; with oligomer arrays, mistaken identity problems and failed PCR amplification are avoided (26,27). To construct a gene-specific microarray, we designed 70mer oligonucleotide immobilized probes using ArrayOligoSelector (27), and the $\sim 10\,000$ open reading frames (ORFs) derived mainly from the *N.crassa* databases at the Broad Institute (http://www.broad.mit.edu/annotation/fungi/neurospora_crassa_7/index.html) and MIPS (http://pedant.gsf.de/cgi-bin/wwwfly.pl?Set=Ncrassa_annotations&Page=index). ArrayOligoSelector selects a unique segment to identify each ORF considering the predicted sequences of all genes. The program avoids selecting internal self-annealing structures and internal repeat sequences, and preferentially chooses oligonucleotides within a narrow range of GC content, which are biased toward the 3' terminal region of each gene. Oligomers were synthesized (Illumina, San Diego, CA) for 3366 ORFs, which represent $\sim 1/3$ of the predicted full gene set (Supplementary Data S1). A large proportion of the gene set selected was located on linkage groups II and V (manually annotated by MIPS) and included previously characterized genes in *N.crassa*. Analysis of oligomer placement in predicted genes indicated the desired strong 3' bias.

Oligomers were re-suspended in $3\times\ \text{SSC}$ to a final concentration of $40\ \mu\text{M}$ and spotted onto poly(L)-lysine-coated microscopic slides as described previously (9). Aliquots

containing 8 of 40 μM ArrayControl Sense Oligo spots, which were complementary to the eight ArrayControl RNA spikes (Ambion, Austin, TX), were also included. Two spots were printed per slide of each synthesized oligomer.

Strain and culture conditions

To produce conidia, a wild-type laboratory strain, R1-01, was inoculated onto 25 ml solid Vogel's minimal medium (MM) with 1.5% agar and 2% sucrose in a 250 ml flask and grown for 8 days under constant light (~ 300 lux) at 25°C . Growth under constant light conditions was performed to suppress gene expression associated with circadian rhythms (D. Bell-Pederson, personal communication). To harvest conidia free from hyphal fragments, they were suspended in Vogel's MM and passed through eight layers of cheese cloth. The conidial suspension was subsequently diluted to 10^7 conidia/ml in Vogel's MM. The ratio of hyphal fragments to conidia was no more than 1 to 50 by volume. To begin each experiment, 50 ml of the conidial suspension was introduced into each of a series of 250 ml flasks. After 30 min with no agitation at 25°C , one sample was harvested. The remaining flasks containing 50 ml conidial suspensions were incubated at 30°C with constant shaking (200 r.p.m., C25 Incubator-Shaker, New Brunswick Scientific, Edison NJ) and light (~ 300 lux). Subsequent samples were harvested at 1, 2, 4, 8, 12 and 16 h. Samples were harvested by vacuum filtration onto a $0.45\ \mu\text{m}$ nitrocellulose filter (Millipore, Bedford, MA) and immediately frozen in liquid nitrogen. The time 0 sample consists of conidia from an 8-day-old culture that were harvested by adding -80°C ethanol directly to the culture, filtered through cheese cloth and a $0.45\ \mu\text{m}$ nitrocellulose filter, and frozen in liquid nitrogen.

Microscopy

Conidia and hyphae were examined using differential interference contrast optics with a Zeiss Axioskop II microscope. Desiccated and germinating conidia in Vogel's MM (25) were mounted on slides and viewed under a $100\times$ oil immersion objective; 4-, 8- and 16-h-old hyphae were viewed with a $40\times$ oil immersion objective. Micrographs were taken with a Hamamatsu digital camera and captured images were processed using Photoshop software (version 6.0, Adobe).

RNA extraction and cDNA labeling

Frozen samples were ground in liquid nitrogen with a mortar and pestle and total RNA was extracted using TRIzol (Invitrogen Life Technologies, Burlington, ON) following the manufacturer's suggested protocol. A $100\ \mu\text{g}$ sample of total RNA from each sample was further purified using RNeasy Mini Protocol (Qiagen, Valencia, CA). For cDNA synthesis, $20\ \mu\text{g}$ of total RNA was mixed with $5\ \mu\text{g}$ of an anchored 17mer oligo(dT) [oligo(dT)₁₅ VN, where V is any nucleotide except thymidine and N is any nucleotide] and $3.3\ \text{ng}$ of ArrayControl single RNA spike mixture (Ambion) in $15.5\ \mu\text{l}$ H_2O ; the mixture was incubated at 70°C for 10 min. cDNA was synthesized in a final volume of $30\ \mu\text{l}$ with $500\ \mu\text{M}$ each of dATP, dCTP and dGTP, $200\ \mu\text{M}$ of dTTP, $300\ \mu\text{M}$ of aminoallyl-dUTP, $10\ \text{mM}$ DTT and $100\ \text{U}$ Stratascript reverse transcriptase (Stratagene, La Jolla, CA) in $1\times$ reaction buffer. RNA was hydrolyzed by adding $7.5\ \mu\text{l}$ of $0.5\ \text{M}$ NaOH and

$50\ \text{mM}$ EDTA, incubated for 15 min at 65°C and subsequently neutralized by the addition of $37.5\ \mu\text{l}$ of $1\ \text{M}$ HEPES buffer, pH 7.5. The neutralized cDNA was purified using CyScribe GFX Purification Kit (Amersham Biosciences, Piscataway, NJ) as recommended by the manufacturer except that $60\ \mu\text{l}$ of $0.05\ \text{M}$ sodium bicarbonate was used for elution of the cDNA. The eluted cDNA was dried under vacuum and re-suspended in $25\ \mu\text{l}$ H_2O . For conjugation to fluorescent dyes, $10\ \mu\text{l}$ of $0.05\ \text{M}$ sodium bicarbonate was added to the monofunctional NHS-esters of Cy3 or Cy5 (CyDye Post-Labeling Reactive Dye, Amersham Bioscience, Piscataway, NJ) and $5\ \mu\text{l}$ of the dye solution was added to the cDNA solution. Fluorescent dye was coupled in the dark at room temperature for 1 h; unbound dye was deactivated by adding $15\ \mu\text{l}$ of $4\ \text{M}$ hydroxylamine and incubating the mixture for 15 min in the dark. The labeled cDNA was purified with the CyScribe GFX Purification Kit (Amersham) and dried under vacuum.

Hybridization and image acquisition

Slides were pre-hybridized at 42°C for at least 1 h in a solution containing 50% formamide, $5\times$ SSC, 0.1% SDS and 0.1 mg/ml BSA. The labeled cDNA was re-suspended in $28\ \mu\text{l}$ of hybridization buffer containing 50% formamide, $5\times$ SSC, 0.1% SDS, 0.1 mg/ml BSA, 0.1 mg/ml salmon sperm DNA and heated at 95°C for 3 min before it was pipetted into the space between a microarray slide and a LifterSlip cover glass (Erie Scientific, Portsmouth, NH). Hybridization was carried out for 16 h at 42°C and unbound DNA was washed off as described previously (28). An Axon GenePix 4000B scanner (Axon Instruments, CA) was used to acquire images, and GenePix Pro 4.1 software was used to quantify hybridization signals. Bad spots were flagged automatically by GenePix software and subsequently each slide was inspected manually.

Data analysis

Hybridized spots with at least one of the mean fluorescence intensities for Cy3 or Cy5 that were brighter than mean background intensity plus three standard deviations of background intensity were scored for further analysis if $<2\%$ of pixels were saturated. If both of the duplicated spots met these criteria, only the brighter spot was taken for analysis. This procedure was followed to avoid pseudoreplication, as duplicate spots on the microarray were highly correlated and did not constitute independent data [c.f. (29)]. Signal intensities from the eight spiked positive control cRNAs (Ambion) were used to perform linear normalization of the data. Normalized ratio data were then analyzed using Bayesian Analysis of Gene Expression Levels (BAGEL) software, with which we inferred a relative gene expression level and credible interval for each gene in each experimental time point (30). Estimated gene expression levels were clustered (31) using Hierarchical Clustering Explorer 2.0 (32), in which similarity in expression patterns between genes is measured as Pearson's correlation coefficient and the closest two genes or clusters are successively joined. Distances between clusters represent the average distances between genes in the clusters.

The Functional Catalogue (FunCat) created by MIPS (33,34) and tables from a *N. crassa* genome review (4) were used to group genes according to their cellular or molecular

functions. Over- or under-representation of gene groups across expression gene clusters was evaluated against an expected hypergeometric distribution using the Fisher test function in the statistical software R 1.9 (<http://bioconductor.org>).

Consensus motif searches

Motif searches were conducted using three programs BioProspector (35), MDscan (36) and MEME (37) on segments 500 bp upstream of predicted translational start sites, which were downloaded from the Broad Institute *N.crassa* database release version 3 (http://www.broad.mit.edu/annotation/fungi/neurospora_crassa_3/). Predicted motifs were subsequently analyzed manually. To search for examples of candidate motifs, segments 1000 bp upstream were searched for sequences that matched either the consensus motif site or degenerate sites by using a string search with a PERL script. Enrichment of motifs ($P < 0.001$) was assessed using Fisher's exact test performed by the Fisher test function implemented in the R 1.9 program (<http://bioconductor.org>). A motif logo illustrating the consensus sequence was obtained using WebLogo program (<http://weblogo.berkeley.edu>).

Cross-species analyses of expression profiles

Bidirectional BLAST searches (blastx and tblastn) were performed between the *N.crassa* gene set and the predicted gene set of *U.maydis* (http://www.broad.mit.edu/annotation/fungi/ustilago_maydis/) and between the *N.crassa* gene set and the *D.discoideum* gene set (<http://dictybase.org/>) (38,39). A pair of protein coding genes was judged to be orthologous if each was each other's best bidirectional blast hit with values $<10^{-5}$.

Expression profiles of *N.crassa* genes were cross-examined with gene orthologs belonging to expression gene groups obtained during teliospore germination in *U.maydis* (40) or those obtained during spore germination in *D.discoideum* (41). Over- or under-representation of expression profile groups across expression gene clusters was evaluated against an expected hypergeometric distribution using the Fisher test function in the statistical software R 1.9 (<http://bioconductor.org>).

RESULTS

Experimental design for conidial germination transcriptional profiling

Circuit design for microarray comparisons has been strongly endorsed by statisticians (29,42,43) and has demonstrated dramatically improved resolution in identifying differential gene regulation compared with designs using a universal reference (44–48). The experimental design for transcriptional profiling of conidial germination is shown in Figure 1A. Each arrow represents one hybridization and the arrowhead points to Cy5-labeled cDNA. In this circuit of experimental comparisons, each sample was compared head-to-head with other samples, in a circular, and in some cases, multiple-pairwise fashion. This design maximized the ability to detect differences in expression because the comparisons are between transcripts from each of the developmental stages of greatest interest rather than between the stages and a reference sample.

RNA was isolated from eight time points during the germination process, time 0, time 30 min and times 1, 2, 4, 8, 12

and 16 h post-inoculation. Our microscopic observations on morphological aspects of conidial germination in *N.crassa* were similar to those reported previously [reviewed in (13)]. Briefly, dehydrated conidia harvested from time 0 showed irregular shapes (Figure 1A). When suspended in Vogel's MM, swelling of conidia was observed. For the first 2 h post-inoculation, microscopic changes were not observed, other than swelling. Germ tube emergence was first observed in a small fraction (~2%) of conidia ~2 h after hydration (Figure 1A). At 4 h post-inoculation, germ tubes had formed in the majority (~98%) of conidia (Figure 1B). Numerous germling fusions were observed at this time point (Figure 1A). At 8 h, germ tubes had extended, branched and began to form a mycelial mat that is characteristic of an exponentially growing fungal colony (see 8 and 16 h time points; Figure 1A).

Dormant conidia have large differences in mRNA abundance compared with germinating conidia, which affected the choice of methods for data normalization

Sachs and Yanofsky (49) observed that the poly(A) mRNA fraction of total RNA increased during conidial germination. Consistent with this observation, competitive hybridization between cDNA from dormant conidia (0 h) and that after 1 h of germination showed a skewed distribution of transcription levels toward 1 h (Figure 2A). These data indicated that most (but not all) transcripts identified in the 1 h culture were much more abundant than in dormant conidia. Those few mRNAs at a measurable concentration in dormant conidia were often found at very high levels (Figure 2A). Between 1 and 2 h the data were dispersed diagonally, indicating that the overall difference in the transcriptional profiles of germinating conidia between the 1 and 2 h time points was small (Figure 2B).

To evaluate the ratio of mRNA from each of the 28 comparative hybridizations in this study (Figure 1A), we used control RNA spikes as internal standards. The control spikes consist of eight polyadenylated bacterial mRNAs at concentrations ranging from 50 to 1000 pg/μl, which are complementary to eight ArrayControl Sense oligonucleotides (Ambion). The ArrayControl oligonucleotides were added as duplicate spots to the *N.crassa* oligonucleotide microarrays. A total of 3.3 ng of each control mRNA spike was added to each of the 20 μg total RNA samples for each time point. The relative amount of total mRNA at each time point was then inferred using BAGEL software (Figure 2C). Although there was a 2-fold discrepancy between the poly(A) assay conducted by Sachs and Yanofsky (49) and control spike estimates in this study, both portray the mRNA content at its nadir in dormant conidia and at its peak between 3 and 4 h after hydration, with a subsequent gradual decline during the experimental time period. Because mRNA quantity changes so drastically in the early stages of germination, and because it forms such a small fraction of total RNA in dormant conidia, we used control RNA spikes to normalize the data. The other method of normalization, global normalization, assumes a constant number of transcripts per cell and is inappropriate when the number of transcripts per cell varies dramatically between samples.

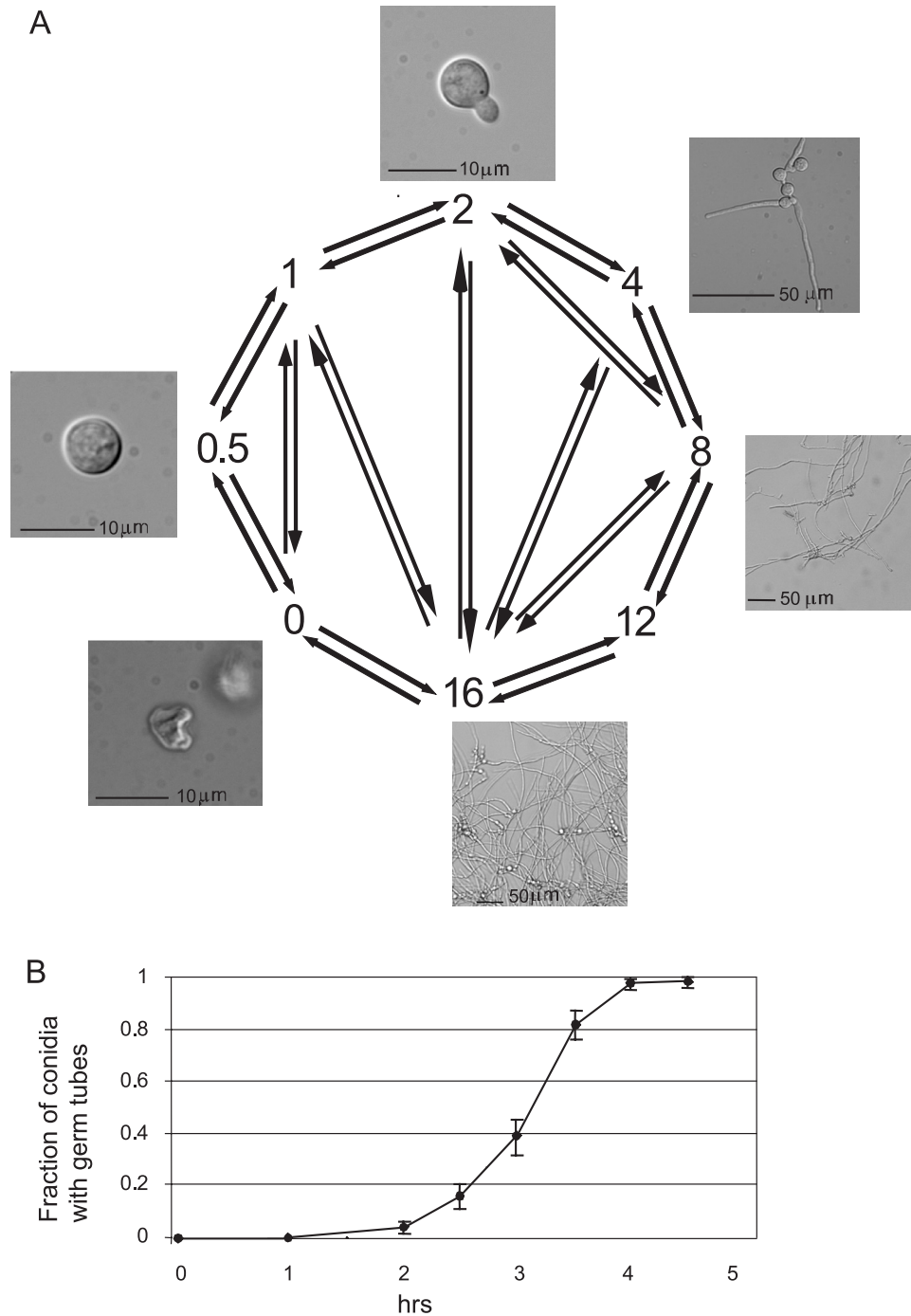


Figure 1. (A) The closed loop experimental design used for transcriptional profiling during conidial germination. Each arrow represents one competitive hybridization and the arrowhead points to Cy5-labeled cDNA. Micrographs of dormant conidia (0) and of cultures 0.5, 2, 4, 8 and 16 h after hydration are shown in panels. (B) Fraction of conidia with germ tubes. For each time point, 200 spores were scored for presence or absence of germ tubes. Confidence of 95% intervals is shown.

Bayesian analysis of the microarray expression data

From the normalized ratios of Cy3 and Cy5 intensities, microarray expression data were analyzed using BAGEL (30). BAGEL uses information from both direct and transitive comparisons to infer relative gene expression levels and 95% confidence intervals across the replicated, interconnected experimental design. The lowest level of expression for a particular gene in the interconnected data is set as a unit

of one. Expression levels of a particular gene in all other samples are scaled appropriately; all gene expression level measurements are positive. The inferred expression levels are of arbitrary unit scale, rather than in absolute counts of mRNA abundance per cell because of the inherent comparative nature of the two-color spotted microarray technology.

The dramatic difference in gene expression between dormant and germinated conidia (0, 0.5 and 1 h) makes the

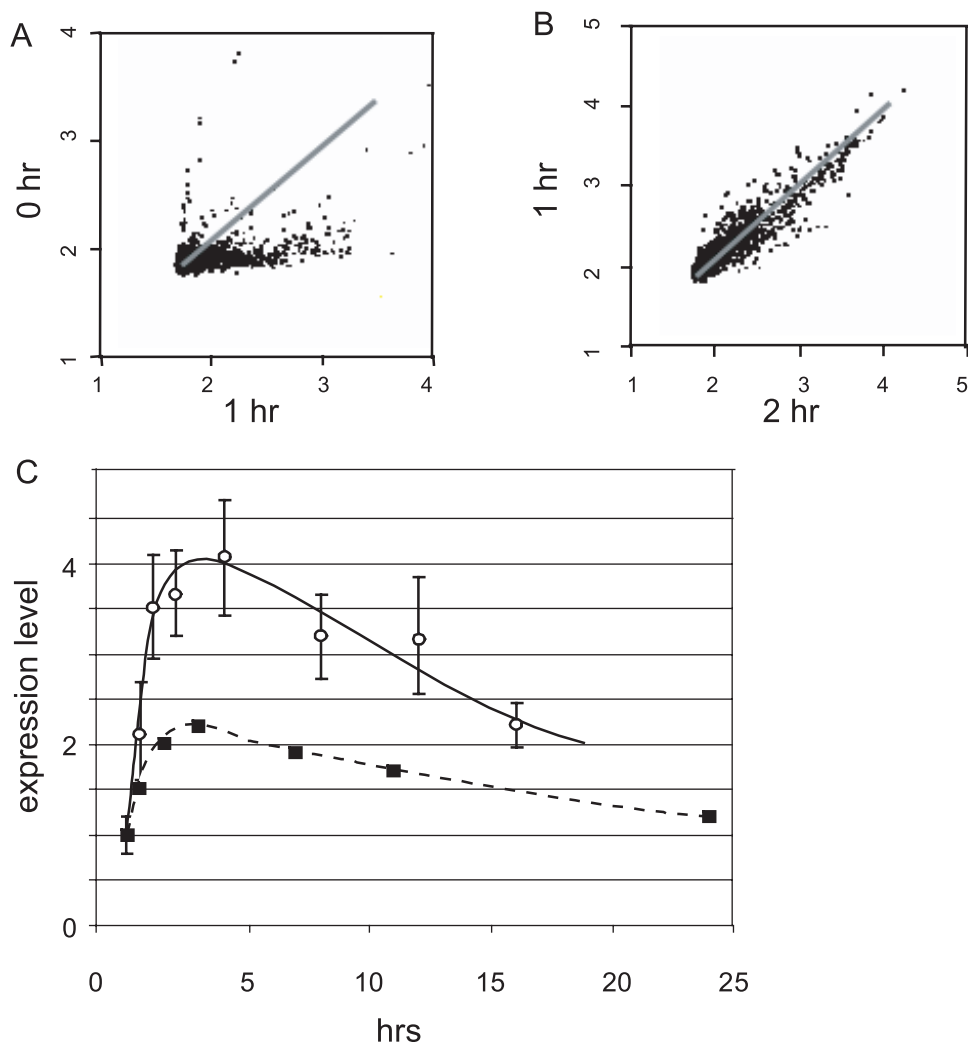


Figure 2. (A) Scatter plot of competitive hybridizations. The median pixel intensities of Cy3-labeled cDNA at 0 h was plotted against Cy5-labeled cDNA from 1 h. (B) The median pixel intensities for Cy3-labeled cDNA from 1 h was plotted against Cy5-labeled cDNA from 2 h. Diagonal lines were drawn over probes for the eight bacterial spike controls. (C) A comparison of changes in mRNA quantity during conidial germination at fixed amount of total RNA is shown. Filled squares: mRNA quantified by poly(A) protection assay [data from (49)]. Open circles: mRNA quantity estimated from microarray data by use of cRNA control spikes and BAGEL software. For both experiments, mRNA quantities were smallest in the dormant conidia, and were set to 1 in relation to other time points.

estimation of gene expression level across all time points difficult because genes expressed at near background will not provide measurements judged significant for subsequent analysis. With circuit designs (Figure 1A), BAGEL requires an interlinked measurement (arrow) for each comparison between two time points, i.e. nine interlinked measurements for eight interconnected time points. We detected hybridization to 3054 (out of a total of 3366) oligonucleotides in the raw data of at least one microarray comparison. Hybridization to 1061 oligonucleotides was detected in a minimum of nine interlinked measurements. Statistical support for relative expression level could not be established for genes that were insufficiently well measured at a number of time points during conidial germination. To use BAGEL to discover other genes with significantly variable relative expression levels, we used smaller circuits that either included only the early time points, or that excluded them. BAGEL analysis of the first small subcircuit, which comprised measurements among time points 0, 0.5, 1, 2 and 16 h (Figure 1A) revealed an

additional 133 genes, many of which were abundant only in conidia. A second small subcircuit composed of measurements among time points 1, 2, 4, 8, 12 and 16 h revealed a further 93 genes, most of which had the lowest expression levels between time 0 and 1 h. Thus, significant data on a total of 1287 genes with relative expression levels and credible intervals were obtained during conidial germination up to 16 h of growth (Supplementary Data S2).

Comparison of microarray data to *N.crassa* expressed sequence tag (EST) databases

Of the 3366 oligos designed for predicted genes, 1748 (52%) have been labeled as 'putative', indicating an EST match, 'hypothetical', indicating no similarity to any protein in the database or 'conserved hypothetical', indicating a similarity to predicted proteins without experimental evidence (7) (Table 1). The genes for these 1748 oligos are termed 'unannotated' in this report, whereas the remaining 1618 oligos, which were

designed to genes with predicted or known gene functions, are termed 'annotated'. Relative gene expression levels for ~50% of the annotated genes (814 out of 1618 genes) were obtained from the interlinked comparisons. In contrast, out of the total of 1748 oligos for unannotated genes, data were obtained for only 27% (473 genes). Unannotated genes for which no transcripts were found may represent genes for which expression is absent, transient or meager during conidial germination, or they may be artifacts of gene prediction algorithms.

EST libraries have previously been constructed and analyzed from germinating conidia, 4.5 h post-inoculation (the conidial library), from 24-h-old mycelium (the mycelial library) and from sexual reproductive structures, 5 days after fertilization (the perithecial library) (50,51). ESTs from a total of 4738 clones were analyzed and assigned to 1721 genetic loci (http://www.broad.mit.edu/annotation/fungi/neurospora_crassa_3/est2loci.html). Of these 1721 loci, 661 have corresponding oligomers on the partial genome microarray (Table 1). A relatively large portion of oligomers to genes on the microarray with matches in the EST libraries had functional annotations (see 'Total ESTs' versus 'Total Oligos' in Figure 3).

Table 1. A summary of oligomers, mRNA profiles and EST matches

| | Oligomers designed | Oligomers yielded mRNA profiles ^a | Oligomers having EST matches ^b |
|---------------|--------------------|--|---|
| Annotated | 1618 | 814 (50.3%) | 452 (27.9%) |
| Not annotated | 1748 | 473 (27.0%) | 209 (11.9%) |
| Total | 3366 | 1287 (38.2%) | 661 (19.6%) |

^amRNA profiles estimated from interlinked experimental design of conidial germination.

^bOligomers having matches to the New Mexico cDNA libraries (50) according to Broad Institute (http://www.broad.mit.edu/annotation/fungi/neurospora_crassa_3/est2loci.html).

Relative expression profiles were obtained for 395 out of the 661 genes on the microarrays with EST matches (69%, Figure 3); 312 of the 395 genes have annotation, while 83 do not. Oligomers for genes with ESTs in the perithecial library showed the lowest percentage of hybridization results, 235 genes out of 413 (57%); the perithecial library is expected to contain ESTs for numerous genes that are specific to perithecial development and function. The mycelial library is expected to contain ESTs for genes that act late in conidial germination, during mycelial growth. Oligomers for genes with ESTs in the mycelial library showed an intermediate percentage of hybridization results, 152 out of 251 genes (61%). The mycelial EST library was constructed from a 24-h-old culture; it is likely that nutrients in the medium are exhausted at this time point. A similar phenomenon in *S.cerevisiae* is associated with large changes in transcriptional profile (52). These observations may explain the lower percentage of hybridization results recovered from the mycelial library. The conidial EST library is expected to contain ESTs for a large proportion of the genes that operate during conidial germination. Indeed, 174 out of 212 genes (82%) that had ESTs in the conidial library yielded hybridization results across the conidial germination time course. Thus, a greater number of oligomers for genes present in a particular EST library yield results when a process is transcriptionally profiled that relates closely to the EST sample conditions. One implication of this observation is that oligomers that lacked hybridization results may not be poorly designed, but rather may correspond to genes that are expressed at detectable levels only in conditions other than those we examined.

Consistency of microarray data with RNA blot analysis

Published northern blot results for several *N.crassa* genes, performed with a fixed amount of total RNA (3 µg), (49) were compared with the microarray results also performed

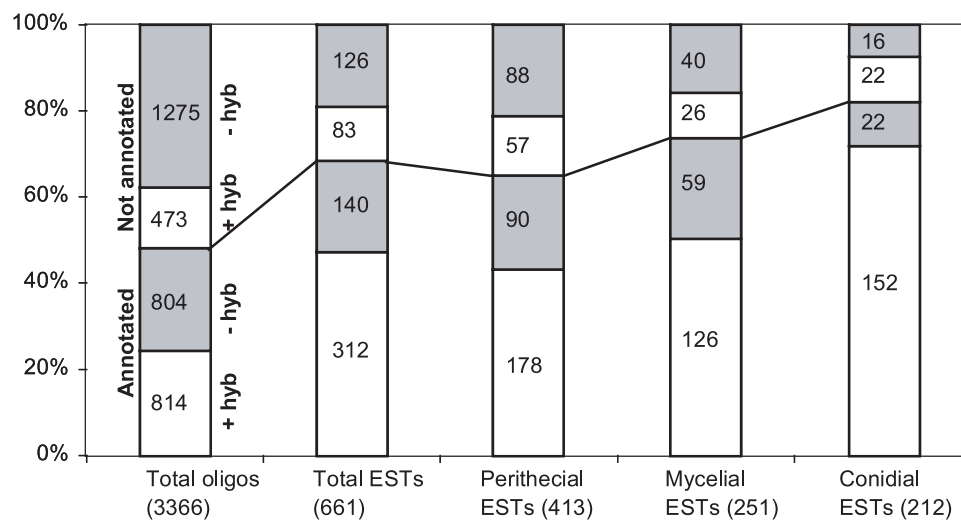


Figure 3. Proportion of genes that yielded mRNA profiles and annotation status of genes to which oligomers were designed and synthesized. From the left, bars are for total genes, all of the predicted genes with EST matches, genes with perithecial EST matches, genes with mycelial EST matches and genes with conidial EST matches. The total number of predicted genes in each category is shown in parentheses. Below the line indicates annotated genes with (open box) or without (shaded box) mRNA profiles. Above the line indicates unannotated genes with (open box) or without (shaded box) mRNA profiles. Number of corresponding genes is shown for each area.

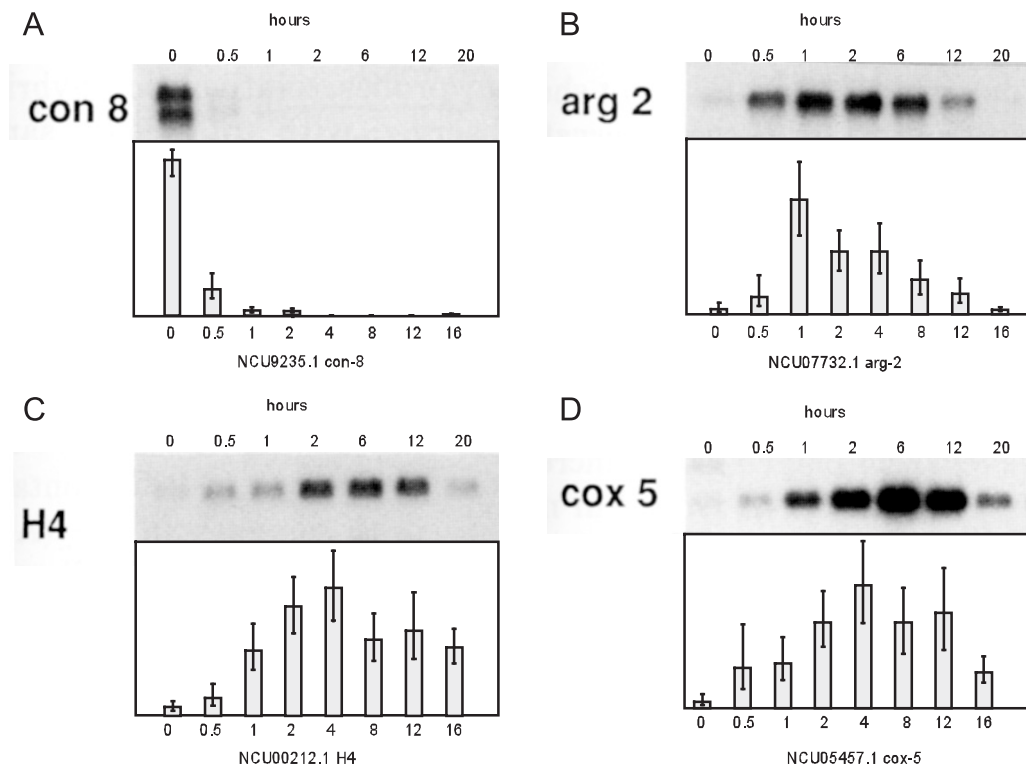


Figure 4. Comparison of northern blot data (49) and microarray data for selected genes that are transcriptional regulated during conidial germination. Three $\mu\text{g}/\text{lane}$ of total RNA was used for the northern data [from (49); used with permission*]. Bars are control cRNA-normalized estimates of relative expression level of mRNA transcripts obtained by BAGEL. An asymmetric 95% credible interval for each expression level is also shown. (A) *con-8*, a conidium-specific gene. (B) *arg-2*, a gene for arginine biosynthesis. (C) *H4*, a histone gene *hH4-1*. (D) *cox-5*, cytochrome-C oxidase chain V gene. Units for the ordinates are arbitrary. *Reprinted from (49) Page 121, Copyright (1991) with permission from Elsevier.

at a fixed amount of total RNA (20 μg). In previous northern blot analyses, transcripts from a conidium-specific gene, *con-8*, were abundant in dormant conidia, but barely detectable at 0.5 h; our *N.crassa* microarray faithfully reproduced this result (Figure 4A). Other conidiation-specific genes on the microarrays including *con-6*, *con-10* (53) and *nop-1* (54) also showed highest expression levels in dormant conidia, with a subsequent decline (Supplementary Data S2), consistent with published northern data. Previously, it was shown that the expression of *arg-2* (a gene for biosynthesis of arginine) increased with the onset of germination (maximum at 1–2 h post-inoculation) and decreased thereafter (49). Again, our microarray results faithfully reproduced these results as compared with technique of northern blotting (Figure 4B). A number of genes that are predicted to encode amino acid biosynthetic enzymes (Figure 6) showed similar expression profiles to that of *arg-2* during conidial germination (Figure 4B). The transcriptional profiles for *hH4-1* (encoding histone H4) and *cox-5* (cytochrome-C oxidase chain V gene) obtained during conidial germination were likewise consistent with published northern data (49) (Figure 4C and D).

Gene clustering of microarray expression profiles during conidial germination

Hierarchical Clustering Explorer (32) was applied to the dataset to cluster genes according to their expression profiles (31) (Figure 5). For those genes whose expression levels were

obtained from the small circuits (time points 0, 0.5, 1, 2 and 16 h or time points 1, 2, 4, 8, 12 and 16 h), it was inferred that the expression level at excluded time points was zero or extremely meager (Supplementary Data S2). For clustering purposes, an expression value of 0 was assigned to the excluded time points. Data were then merged with the larger set of results for genes with full time-course mRNA expression levels. The dendrogram was constructed from distances (1-correlation coefficient) between expression patterns of genes (see Materials and Methods). Seven representative clusters were chosen visually, each of which showed distinctive profiles (Figure 5). The majority of genes (1153 out of the total of 1287 genes) belonged to one of the seven chosen clusters. The genes in the Max0 cluster showed maximum expression levels in dormant conidia, with expression levels dropping quickly during conidial germination. The Max0.5 cluster included genes with a peak expression at 30 min. Two clusters, Max1A and Max1B, which showed maximum expression at 1 h, followed by a decrease in expression levels, clustered in different locations in the dendrogram. Genes within the Max1A and Max1B clusters were pooled and are referred to as Max1 cluster hereafter. Genes within the Max1~4 cluster showed an expression plateau at 1–4 h post-inoculation. The Max8~16 cluster contained genes that had low expression levels until 4 h, but abundant expression from 8 h onward. Genes within the Min0~0.5 cluster showed the lowest expression levels in dormant conidia to 30 min post-inoculation, but higher and relatively constant expression levels thereafter.

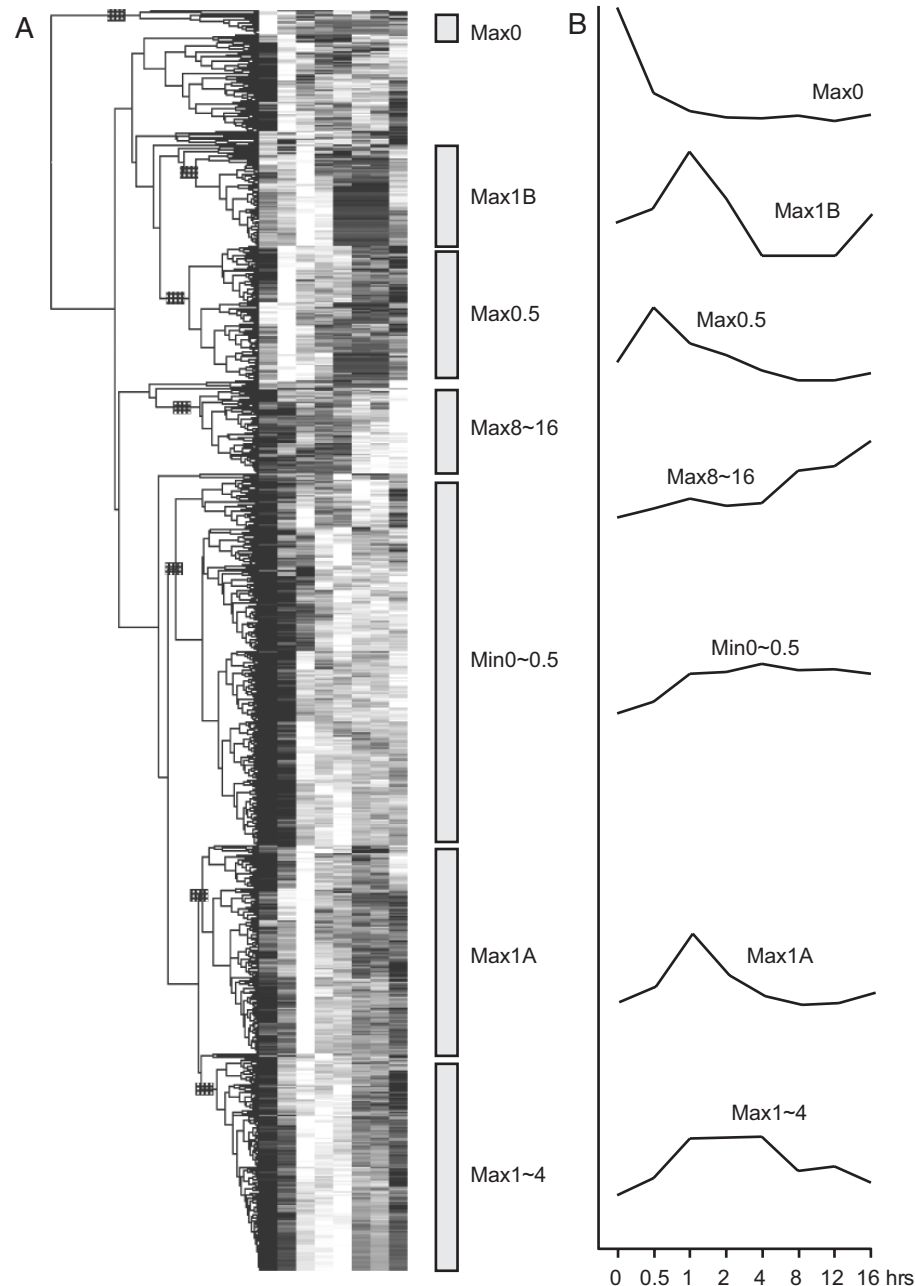


Figure 5. Hierarchical clustering of microarray data and identification of genes with similar transcriptional profiles. (A) A total of 1287 genes were clustered based on their expression profiles across the eight time points. Each gene's expression values were standardized to have mean zero and standard deviation of one across the eight time points. The lighter color in the cluster dendrogram is correlated with a higher expression level. Seven distinct clusters were visually selected, and the node separating each cluster is shown in the distance tree. (B) An average expression profile of genes within each cluster. To obtain each profile, a sum of each gene's expression values across the eight time points was standardized to 1. Next, time course values for all genes in each cluster were summed, and the summed value for the eight time points for each cluster was scaled to 1. The name designation of each cluster and representative profile is shown.

Analysis of clustered genes via FunCat classification revealed major cellular events associated with conidial germination

A recent criticism against the use of mRNA profiling in characterizing cell phenotypes is based on observations in which transcriptional profile does not directly correlate with the proteomic profile, the abundance and activity of the components of which actively determine cell phenotype (11,12). However, the utility of mRNA profiling depends upon its correlation

with and implications for functional measurements of cell biology, biochemistry and organismic development, not upon the bivariate correlation of high-throughput methodologies. We therefore cross-examined mRNA profiling data, gene function and cellular events in order to validate mRNA profiling.

The MIPS FunCat is an annotation scheme for the functional description of proteins (33,34). FunCat consists of 28 main functional categories, each of which contains

hierarchically ordered subcategories with increasing levels of specificity. A single gene can belong to more than one functional category. Out of the 1618 genes with functional annotation which were included in the *N.crassa* oligonucleotide microarrays, 814 had time-course mRNA profile data (Table 1). The main functional categories populated by genes from each of the six clusters are summarized in Table 2.

Genes predicted to be in a particular functional category were not distributed evenly among the six clusters. For example, genes with transcriptional function were over-represented in Max1 and Max1~4 clusters, but under-represented in Max0, Max8~16 and Min0~0.5 clusters ($P = 0.009$) (Table 2). Subsets of FunCat 'Transcription' were evaluated for their impact on the uneven distribution of genes within the clusters. Genes that function in RNA synthesis were particularly over-represented in the clusters Max1 and Max1~4 ($P = 0.021$). These data are consistent with biochemical data indicating the RNA synthesis is activated soon after the induction of germination (13).

Genes that function in the cell cycle and DNA processing were over-represented in the Max0.5, Max1, Max1~4 clusters, but were under-represented in Max0 and Min0~0.5 clusters ($P = 0.013$). A subset of these genes putatively involved in DNA processing were over-represented in clusters Max0.5 and Max1~4 ($P = 0.039$). These data are consistent with biochemical observations that the initiation of DNA replication occurs approximately 2 h post-inoculation (13), just before germ tube formation (between 1–2 h; Figure 1A).

Genes that function in protein synthesis were over-represented in Max1~4 cluster ($P < 0.001$), as were genes within the 'Subcellular localization' MIPS category ($P = 0.000$). A subset of protein synthesis genes that function in ribosome biogenesis deviated remarkably from an even distribution among the gene clusters ($P = 0.000$) (Table 2). In fact, 49 out of 56 genes functioning in ribosome biogenesis, and all 24 of those that have been identified as ribosomal proteins, were members of the cluster Max1~4 and showed strikingly similar expression patterns (Figure 6). The majority of genes categorized as functioning in ribosome biogenesis are also categorized within the 'Subcellular Localization' category, resulting in genes belonging to this category also being over-represented in Max1~4. Ribosomes and mRNA are stored in dormant conidia and protein biosynthesis is detectable within the first few minutes after the conidia have been suspended in water, salts and a carbon source (13). A large fraction of macromolecular synthesis in germinating conidia is devoted to ribosomal RNA and the protein synthetic machinery (2,49). These data on increase in mRNA levels of genes encoding ribosomal protein genes are also consistent with previously published northern and biochemical data (13,49).

Transcriptional profiles for 17 heat shock proteins (HSPs) (4) were identified during conidial germination. Thirteen out of seventeen HSPs showed the highest level of expression between 0.5 and 1 h after the rehydration of conidia (Figure 6); eleven of these belong to either the Max0.5 or Max1 cluster. In another study, proteins levels of HSC70 (NCU09602.1, B18E6_040) were evaluated during development; high levels of HSC70 were associated with aerial hyphae formation and conidia (55). These observations have led to the speculation that HSPs may be important in

correct folding of proteins of newly synthesized proteins during conidial germination. Our transcriptional profiling data supports such a hypothesis.

Biochemical analysis of amino acid pools in conidia and germinating conidia showed that conidial extracts contain detectable free pools of amino acids, with the exception of proline, methionine and cysteine. The amino acid pools present in conidia decrease rapidly at the initiation of germination, but rapidly increase ~3 h post-inoculation, including arginine, ornithine and citrulline pools (56). Transcripts from 25 genes predicted to be involved in amino acid biosynthesis were detected as differentially expressed during conidial germination. Eighteen of these genes belonged to clusters for very early time points: one in Max0 (NCU01830.1, B23G1_170: probable 4-hydroxyphenylpyruvate dioxygenase), three in Max0.5 (NCU02333.1, 7nc525_140: arginase, NCU03748.1, 93G11_270: probable saccharopine reductase and NCU04856.1, xnc010_060: probable glutamine synthetase) and 14 in Max1.

Genes for 'Cellular transport, transport facilities and transport routes' were over-represented in the Max1~4, Max8~16 and Min0~0.5 clusters, but under-represented in the Max0 and Max0.5 clusters ($P = 0.003$). This set of genes encodes diverse types of proteins, such as ATPases, ATP synthases, small molecule and ion transporters, mitochondrial translocators, cytochrome C-related enzymes and NADH-related enzymes. Subsets of these genes that function in 'Transport facilities' ($P = 0.004$) and in 'Transported compounds' ($P = 0.024$) explain the over-representation of genes in the Max1~4, Max8~16 and Min0~0.5 clusters (i.e. vacuolar ATP synthase subunit F; NCU04387.1, 29e8_280). These data are also consistent with previous biochemical data, indicating that rate of transport of nucleosides, glucose and amino acid increases during conidial germination (13,57).

Genes within the 'Energy' functional category were also over-represented in the Max1~4, Max8~16 and Min0~0.5 clusters, but under-represented in Max0, Max0.5 and Max1 clusters ($P < 0.001$). Subsets of these genes functioning in respiration ($P = 0.010$) and electron transport and membrane-associated energy conservation ($P = 0.003$) were partly responsible for this over-representation in Max8~16 and Min0~0.5 clusters. Eight NADH-related enzymes in this category, such as NADH dehydrogenases and NADH-ubiquinone oxidoreductase, showed low gene expression levels at early time points during conidial germination and belong to either Max8~16 or Min0~0.5 cluster. These data are consistent with data showing that conidia have a relatively low rate of oxygen consumption (13,58), which increases greatly 2 to 4 h post-inoculation. More than half of the genes represented on the microarray that function in respiration, electron transport and membrane-associated energy conservation have transcriptional profiles that belong to the Min0~0.5 cluster. The function of *N.crassa* genes involved in the central metabolism (glycolysis, gluconeogenesis, TCA cycle, glyoxylate cycle, fermentation, pentose phosphate pathway) have been deduced from functions of orthologous genes in *S.cerevisiae* [Figure 2 in (59)]. The transcriptional profiles of 16 genes involved in central metabolism were identified in this study; most belonged to the Min0~0.5 cluster.

Table 2. Distribution of genes belonging to specific functional categories among the six hierarchical clusters

| Main category according to FunCat or (4) | FunCat sub category ^a or table number | Observed or expected ^b | Max0 | Max0.5 | Max1 | Max1~4 | Max8~16 | Min0~0.5 | Other | Total ^c | P-value ^d |
|---|--|-----------------------------------|------|--------|------|--------|---------|----------|-------|--------------------|----------------------|
| Biogenesis of cellular components | | Obs | 0 | 4 | 6 | 5 | 1 | 7 | 1 | 24 | 0.914 |
| | | Exp | 2.6 | 3.4 | 4.3 | 2.6 | 1.7 | 8.6 | 0.9 | | |
| Cell cycle and DNA processing | | Obs | 0 | 16 | 16 | 17 | 1 | 16 | 6 | 72 | 0.013 |
| | | Exp | 7.7 | 10.3 | 12.9 | 7.7 | 5.1 | 25.7 | 2.6 | | |
| Cell cycle and DNA processing | DNA processing (2nd) | Obs | 0 | 7 | 4 | 8 | 0 | 4 | 2 | 25 | 0.039 |
| | | Exp | 0.5 | 2.7 | 6.1 | 4.2 | 1.7 | 7.2 | 2.6 | | |
| Cell fate | | Obs | 3 | 6 | 11 | 7 | 1 | 15 | 4 | 47 | 0.403 |
| | | Exp | 5.0 | 6.7 | 8.4 | 5.0 | 3.4 | 16.8 | 1.7 | | |
| Cell rescue, defense and virulence | | Obs | 2 | 8 | 12 | 6 | 7 | 18 | 4 | 57 | 0.307 |
| | | Exp | 6.1 | 8.1 | 10.2 | 6.1 | 4.1 | 20.4 | 2.0 | | |
| Cell type differentiation | | Obs | 3 | 8 | 10 | 6 | 2 | 12 | 3 | 44 | 0.247 |
| | | Exp | 4.7 | 6.3 | 7.9 | 4.7 | 3.1 | 15.7 | 1.6 | | |
| Cellular communication/signal transduction mechanism | | Obs | 0 | 1 | 13 | 6 | 2 | 22 | 4 | 48 | 0.128 |
| | | Exp | 5.1 | 6.9 | 8.6 | 5.1 | 3.4 | 17.1 | 1.7 | | |
| Cellular transport, transport facilities and transport routes | | Obs | 0 | 7 | 22 | 23 | 11 | 44 | 4 | 111 | 0.003 |
| | | Exp | 11.9 | 15.9 | 19.8 | 11.9 | 7.9 | 39.6 | 4.0 | | |
| Cellular transport, transport facilities and transport routes | Transport facilities (2nd) | Obs | 0 | 1 | 8 | 12 | 8 | 22 | 2 | 53 | 0.004 |
| | | Exp | 1.1 | 5.7 | 12.9 | 8.9 | 3.5 | 15.4 | 5.5 | | |
| Cellular transport, transport facilities and transport routes | Transported compounds (2nd) | Obs | 0 | 3 | 11 | 14 | 9 | 19 | 2 | 58 | 0.024 |
| | | Exp | 1.2 | 6.2 | 14.1 | 9.8 | 3.9 | 16.8 | 6.0 | | |
| Classification not yet clear-cut | | Obs | 1 | 3 | 14 | 2 | 3 | 10 | 6 | 39 | 0.210 |
| | | Exp | 4.2 | 5.6 | 7.0 | 4.2 | 2.8 | 13.9 | 1.4 | | |
| Energy | | Obs | 0 | 1 | 5 | 9 | 11 | 28 | 5 | 59 | 0.000 |
| | | Exp | 6.3 | 8.4 | 10.5 | 6.3 | 4.2 | 21.1 | 2.1 | | |
| Energy | Respiration (2nd) | Obs | 0 | 0 | 2 | 3 | 4 | 12 | 0 | 21 | 0.010 |
| | | Exp | 0.4 | 2.3 | 5.1 | 3.5 | 1.4 | 6.1 | 2.2 | | |
| Energy | Electron transport and membrane-associated energy conservation (2nd) | Obs | 0 | 0 | 1 | 3 | 5 | 10 | 0 | 19 | 0.003 |
| | | Exp | 0.4 | 2.0 | 4.6 | 3.2 | 1.3 | 5.5 | 2.0 | | |
| Interaction with the cellular environment | | Obs | 2 | 5 | 15 | 8 | 4 | 14 | 1 | 49 | 0.323 |
| | | Exp | 5.3 | 7.0 | 8.8 | 5.3 | 3.5 | 17.5 | 1.8 | | |
| Protein fate (folding, modification and destination) | | Obs | 2 | 11 | 28 | 13 | 6 | 38 | 11 | 109 | 0.758 |
| | | Exp | 11.7 | 15.6 | 19.5 | 11.7 | 7.8 | 38.9 | 3.9 | | |
| Protein synthesis | | Obs | 0 | 2 | 15 | 60 | 0 | 6 | 2 | 85 | 0.000 ^f |
| | | Exp | 9.1 | 12.1 | 15.2 | 9.1 | 6.1 | 30.4 | 3.0 | | |
| Protein synthesis | Ribosome biogenesis (2nd) | Obs | 0 | 1 | 6 | 49 | 0 | 0 | 0 | 56 | 0.000 |
| | | Exp | 1.1 | 6.0 | 13.6 | 9.4 | 3.7 | 16.2 | 5.8 | | |
| Protein synthesis | Ribosomal protein (3rd) | Obs | 0 | 0 | 0 | 24 | 0 | 0 | 0 | 24 | 0.000 |
| | | Exp | 0.5 | 2.6 | 5.8 | 4.0 | 1.6 | 7.0 | 2.5 | | |
| Protein with binding function or cofactor requirement (structural or catalytic) | | Obs | 0 | 3 | 12 | 8 | 4 | 15 | 0 | 42 | 0.177 |
| | | Exp | 4.5 | 6.0 | 7.5 | 4.5 | 3.0 | 15.0 | 1.5 | | |
| Subcellular localization | | Obs | 1 | 23 | 51 | 71 | 15 | 72 | 16 | 249 | 0.000 ^f |
| | | Exp | 26.7 | 35.6 | 44.5 | 26.7 | 17.8 | 88.9 | 8.9 | | |
| Subcellular localization | Mitochondrion (2nd) | Obs | 0 | 5 | 16 | 28 | 8 | 21 | 4 | 82 | 0.002 |
| | | Exp | 0.8 | 4.1 | 9.3 | 6.5 | 2.6 | 11.1 | 4.0 | | |
| Subcellular localization | Cytoplasm (2nd) | Obs | 0 | 3 | 11 | 27 | 2 | 13 | 4 | 60 | 0.000 |
| | | Exp | 0.8 | 4.1 | 9.3 | 6.5 | 2.6 | 11.1 | 4.0 | | |
| Metabolism | | Obs | 4 | 19 | 55 | 24 | 14 | 55 | 31 | 202 | 0.105 |
| | | Exp | 21.6 | 28.9 | 36.1 | 21.6 | 14.4 | 72.1 | 7.2 | | |
| Metabolism | Amino acid biosynthesis (2nd) ^e | Obs | 1 | 3 | 14 | 0 | 0 | 4 | 3 | 25 | 0.004 |
| | | Exp | 0.5 | 2.7 | 6.1 | 4.2 | 1.7 | 7.2 | 2.6 | | |
| Transcription | | Obs | 1 | 13 | 34 | 18 | 1 | 25 | 4 | 96 | 0.009 |
| | | Exp | 10.3 | 13.7 | 17.1 | 10.3 | 6.9 | 34.3 | 3.4 | | |
| Transcription | RNA synthesis (2nd) | Obs | 1 | 13 | 31 | 18 | 1 | 25 | 4 | 93 | 0.021 |
| | | Exp | 1.9 | 10.0 | 22.6 | 15.7 | 6.2 | 27.0 | 9.7 | | |
| Unclassified proteins | | Obs | 12 | 55 | 108 | 60 | 38 | 134 | 66 | 473 | 0.002 ^f |
| | | Exp | 50.7 | 67.6 | 84.5 | 50.7 | 33.8 | 168.9 | 16.9 | | |
| Transcription factors | Tables 9–14 in (4) | Obs | 1 | 5 | 9 | 1 | 1 | 9 | 1 | 27 | 0.017 |
| | | Exp | 0.5 | 2.9 | 6.6 | 4.6 | 1.8 | 7.8 | 2.8 | | |
| Heat shock proteins | Table 50 in (4) | Obs | 1 | 7 | 4 | 3 | 0 | 2 | 0 | 17 | 0.007 |
| | | Exp | 0.3 | 1.8 | 4.1 | 2.9 | 1.1 | 4.9 | 1.8 | | |
| Glucan and chitin synthesis | Tables 51–52 in (4) | Obs | 0 | 2 | 2 | 1 | 0 | 4 | 1 | 10 | 0.913 |
| | | Exp | 0.2 | 1.0 | 2.3 | 1.6 | 0.6 | 2.8 | 1.0 | | |
| | Total ^g | | 26 | 138 | 313 | 217 | 86 | 373 | 134 | 1287 | |

^aOrdinal numbers in the parentheses indicate levels of FunCat subcategories (33).^bObserved number of genes and expected number of genes if probabilities of each outcome are independent of the cluster.^cNumber of genes in each of the subcategories.^dHypergeometric probability was used to obtain the *P*-value using fisher.test() in R (bioconductor.org).^eGenes involving in amino acid biosynthesis were selected from a MIPS subcategory 'amino acid metabolism (2nd)'.^fChi-square test was used to obtain the *P*-value using chisq.test() in R (bioconductor.org).^gNumber of genes in each of the hierarchical clusters.

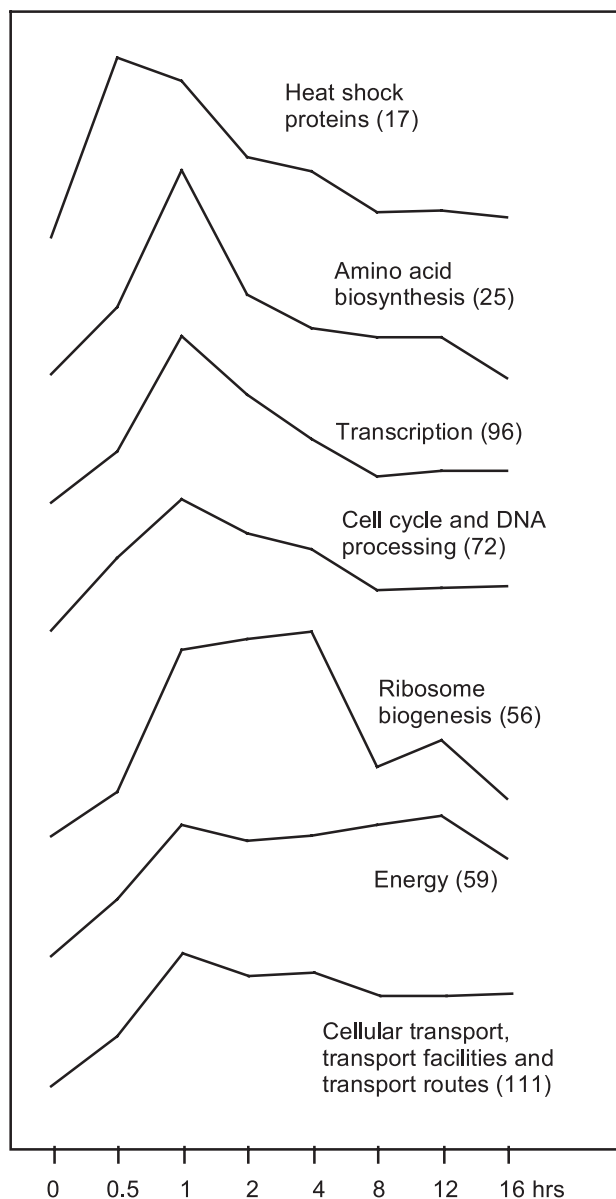


Figure 6. The re-scaled expression profiles of genes from specified functional categories within a cluster. The averaged profile contours were obtained as in Figure 5B for each group of genes.

Given that *Neurospora* hyphae contain chitin (~1%) and abundant glucans and that conidia and conidial germ tubes and mature hyphae have different amounts of cell wall components (60,61), it is not surprising that transcriptional profiles of 10 genes predicted to be involved in glucan and chitin synthesis [Tables 51 and 52 in (4)] were identified during conidial germination. All of these genes have minimal expression in dormant conidia and reached peak expression levels at 0.5 h and afterwards.

Enrichment of FunCat categories in clusters enabled us to infer participation of genetic process during conidial germination. Correlation between gene expression and quantity of active protein during conidial germination notwithstanding, it was demonstrated that mRNA expression and major cellular events were highly associated.



Figure 7. Consensus sequence motif found in the promoter region of 34 putative ribosomal protein genes from the Max1~4 cluster using BioProspector program (65) and WebLogo (<http://weblogo.berkeley.edu>).

Transcripts from hypothetical genes

Genes coding for unclassified proteins are described as 'hypothetical', 'conserved hypothetical' or 'putative proteins' (6,7). Transcriptional profiles for 473 unclassified genes were identified during conidial germination by the *N.crassa* microarrays. Of these 473 genes, 152 are listed as hypothetical, 252 are listed as conserved hypothetical and 69 are putative genes. These unannotated genes were over-represented in Max1 and Max1~4, but otherwise were found abundantly in all the clusters (Table 2). For example, a hypothetical gene, NCU09614.1, showed maximum expression 30 min after the initiation of germination, followed by a steep decrease in expression levels, a profile identical to a number of the heat shock genes (Figure 6). The expression pattern of another putative gene, NCU05338.1 (B15B24_010), belongs to Max8~16; its transcripts were almost undetectable until 8 h, but showed high expression at 8, 12 and 16 h. The Max8~16 cluster contains many genes that function in fermentation and carbohydrate metabolism.

Promoter analysis of clustered genes reveals putative regulatory sequences

Transcriptional profiling has been useful in identifying co-regulated genes and predicting both function and consensus binding sites for transcription factors (35,62–64). To find such nucleotide motifs, we evaluated genes within five clusters, Max0, Max0.5, Max1, Max1~4 and Max8~16, using three programs, BioProspector (65), MDscan (36) and MEME (37). We focused on genes within each cluster that showed defined expression profiles with statistical support. For most of the clusters, a statistically significant consensus sequence motif was not identified, probably because clusters based on similarity of expression pattern contained genes belonging to diverse functional categories with distinct transcriptional regulatory mechanisms. However, when 36 representative genes belonging to Max1~4 were examined, a consensus motif (G/A)CCC(T/C)AA was identified in 26 genes (Figure 7). Subsequent analysis of these 26 genes showed that all were annotated as ribosomal or probable ribosomal protein genes (Table 3). A very similar sequence was previously identified by a purely computational approach that assessed putative *cis*-regulatory sequences in ribosomal protein genes in a number of eukaryotic species (*N.crassa*, *M.grisea* and *Aspergillus nidulans* [Figure 3 in (62)]). This *cis*-regulatory sequence is not found in the promoter region of ribosomal protein genes in *S.cerevisiae*. The entire Max1~4 cluster from *N.crassa* contains 211 genes. Of these 211 genes, 48 genes were annotated as (probable) ribosomal protein

Table 3. Motif enrichment analysis

| Motif | Selected genes from Max1~4 cluster ^{a,b} | Ribosomal proteins in Max1~4 cluster ^{a,c} | Total of 1257 genes with mRNA profiles |
|-------------|---|---|--|
| GCCCTAA | 14 (3.84×10^{-9}) | 16 (1.07×10^{-9}) | 55 |
| G/CCCCT/CAA | 26 (4.17×10^{-3}) | 34 (8.85×10^{-4}) | 414 |

P-values for enrichment of strict and degenerated motifs according to fisher's exact test were shown in parentheses.

^aThere were 211 genes that belong to Max1~4 cluster.

^bThirty-seven genes were evaluated (see Materials and Methods).

^cForty-eight genes are annotated as (putative) ribosomal protein genes out of 211 genes in the Max1~4 cluster.

genes. The Max1~4 cluster was searched for genes containing the (G/AC)CC(T/C)AA motif within 1000 bp of the predicted translational start site and an additional eight genes annotated as (probable) ribosomal protein genes were recovered (Table 3). Thus, out of the 48 ribosomal protein genes in *N.crassa* that showed peak expression 1–4 h post-inoculation, 34 have a consensus motif in their promoter region that may be involved in their transcriptional regulation during conidial germination. Coordinate expression of ribosomal protein genes that respond to altered nutritional conditions have also been reported (66–68). Two consensus sequences in the promoter regions of *crp-2* gene [encoding 40S ribosomal protein S17 (66)] were identified by deletion analysis of the *crp-2* promoter region (68). The (G/A)CCC(T/C)AA consensus motif in the promoter of *crp-2* is next to, but does not overlap, the –74 to –66 bp CG repeat, which was shown to be critical for transcriptional regulation based on carbon source (67).

The remaining genes within the Max1~4 cluster did not contain the (G/A)CCC(T/C)AA consensus motif (Table 3), indicating that other regulatory mechanisms give a similar transcriptional profile to their target genes as those genes involved in ribosome biogenesis. Interestingly, NCU02075.1 (1nc340_120) and NCU03009.1 (B24P7_270) are putative homologs of *S.cerevisiae* heat shock proteins Ssb and Zuo (4); Ssb and Zuo constitute the *S.cerevisiae* ribosomal-associated chaperone machinery (69–71). Similar to the putative ribosomal protein genes in Max1~4 cluster, NCU02075.1 (1nc340_120) and NCU03009.1 (B24P7_270) showed peak expression levels between 1 and 4 h, although the (G/A)CCC(T/C)AA consensus motif was not identified in their promoter region.

Comparison of conidial germination transcriptional profiles with transcriptional profiles of germinating spores of *U.maydis* and *D.discoideum*

The ability to persist through harsh conditions and emerge from a dormant state, as is manifest in spore germination of bacteria, protozoans, ferns and fungi and seed germination of flowering plants, may rely on some universal genetic properties. Recently, a number of cDNA microarray studies have identified differentially expressed genes during spore germination in a basidiomycete fungus *U.maydis* (40), in a water fern *Ceratopteris richardii* (72), in the social amoeba *D.discoideum* (41) and during seed germination in *Brassica oleracea* (73). Among the diverse organisms above, the genome

sequence and cDNA microarray data for *U.maydis* and *D.discoideum* were the most complete. We performed a cross-species comparison of mRNA profiles during spore germination between *N.crassa* and *U.maydis* and between *N.crassa* and *D.discoideum* to evaluate the hypothesis of conservation of processes associated with the breaking of dormancy and initiation of germination among phylogenetically diverse organisms. Of these three organisms, spore germination in *N.crassa* has been most well characterized on a biochemical and physiological level. Bidirectional BLAST searches identified 3117 *U.maydis* orthologs and 2115 *D.discoideum* orthologs to the predicted 10032 *N.crassa* gene set. Expression profiles of these orthologous genes during germination were then cross-examined.

U.maydis is a phytopathogenic fungus and produces dormant propagules called teliospores on maize. Teliospores are diploid, and upon germination, undergo meiosis to form haploid basidiospores. These haploid basidiospores subsequently undergo mating with a cell of the opposite mating type to form the infectious dikaryotic hyphae. Via hybridization to a cDNA microarray representing ~1/3 of the *U.maydis* genome, Zahiri *et al.* (40) identified a number of genes that were differentially expressed at 4 and 11 h after the induction of germination as compared with dormant teliospores. At 4 h post-inoculation, germ tubes of *U.maydis* start to emerge, a time point morphologically similar to *N.crassa* conidia at 2 h post-inoculation. At an 11 h time point, germ tubes of *U.maydis* are elongated, corresponding to a morphological time point of ~4hr during conidial germination in *N.crassa*. A comparison of *U.maydis* clusters to the *N.crassa* conidial germination profiling dataset showed that orthologs of genes belonging to the '11 h up' cluster in *U.maydis* were not distributed evenly in the *N.crassa* conidial germination clusters (Table 4). These orthologs were over-represented in the Max1~4 cluster ($P = 0.011$). Gene orthologs found in the both Max1~4 and '11 h up' were mainly ribosomal proteins (Supplementary Data S3). Similarly, orthologous genes belonging to the '11 h down' cluster in *U.maydis* were over-represented in the *N.crassa* Max0.5 and Max1 clusters and under-represented in the Max1~4 cluster ($P = 0.000$). Four over-represented gene orthologs, which belonged to Max 0.5, Max1 and '11 h down' showed peak expression at very early time points in both fungi, with expression declining upon germ tube emergence. These genes belonged to diverse FunCat categories (Supplementary Data S3).

The social soil amoeba, *D.discoideum* (41), produces asexual dormant spores in response to starvation. Two morphological transitions are associated with spore germination in *D.discoideum*. First, similar to conidial germination in *N.crassa*, spores swell and lose birefringence, which occurs in the first 1–3 h post-activation. The second transition occurs when spore cases break and amoebae emerge, which begins to occur at 2 h post-activation. Using a cDNA array to 7744 genes, Xu *et al.* (41) identified three expression clusters during spore germination in *D.discoideum*. *N.crassa* is evolutionarily distant from *D.discoideum* and a smaller number of gene orthologs were identified via a cross-species comparison, as compared with the *N.crassa* and *U.maydis* common gene set. Nevertheless, orthologs of genes belonging to the cluster III in *D.discoideum* were over-represented in the *N.crassa* Max1 cluster and under-represented in the Max1~4 cluster

Table 4. Cross-species comparisons of spore-germination profiles comprising gene orthologs

| Expression profile group | Observed or expected ^a | <i>N.crassa</i> conidial germination profiles | | | | | | | Total ^b | <i>P</i> -value ^c |
|--|-----------------------------------|---|--------|------|--------|---------|----------|-------|--------------------|------------------------------|
| | | Max0 | Max0.5 | Max1 | Max1~4 | Max8~16 | Min0~0.5 | Other | | |
| <i>U.maydis</i> teliospore germination (40) ^d | | | | | | | | | | |
| 4 H up | Obs | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0.560 |
| | Exp | 0.0 | 0.1 | 0.1 | 0.4 | 0.2 | 0.2 | 0.1 | | |
| 4 H down | Obs | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1.000 |
| | Exp | 0.0 | 0.1 | 0.1 | 0.4 | 0.2 | 0.2 | 0.1 | | |
| 11 H up | Obs | 0 | 0 | 0 | 9 | 1 | 3 | 0 | 13 | 0.011 |
| | Exp | 0.0 | 1.0 | 1.0 | 5.2 | 2.1 | 2.6 | 1.0 | | |
| 11 H down | Obs | 0 | 2 | 2 | 1 | 3 | 0 | 2 | 10 | 0.000 |
| | Exp | 0.0 | 0.8 | 0.8 | 4.0 | 1.6 | 2.0 | 0.8 | | |
| Total ^f | | 0 | 2 | 2 | 10 | 4 | 5 | 2 | 25 | |
| <i>D.discoideum</i> spore germination (41) ^e | | | | | | | | | | |
| Cluster I (Max in dormant spores) | Obs | 0 | 1 | 0 | 1 | 0 | 3 | 0 | 5 | 0.769 |
| | Exp | 0.0 | 0.4 | 0.4 | 2.0 | 0.8 | 1.0 | 0.4 | | |
| Cluster II (Max at 6 h) | Obs | 0 | 1 | 0 | 1 | 0 | 3 | 0 | 5 | 0.769 |
| | Exp | 0.0 | 0.4 | 0.4 | 2.0 | 0.8 | 1.0 | 0.4 | | |
| Cluster III (Max at 2 h) | Obs | 0 | 0 | 3 | 0 | 1 | 2 | 0 | 6 | 0.030 |
| | Exp | 0.0 | 0.5 | 0.5 | 2.4 | 1.0 | 1.2 | 0.5 | | |
| Total ^f | | 0 | 2 | 3 | 2 | 1 | 8 | 0 | 16 | |

^aObserved number of genes and expected number of genes if probabilities of each outcome are independent of the cluster.^bNumber of gene orthologs in each of the expression profile groups.^cHypergeometric probability was used to obtain the *P*-value using `fisher.test()` in R (bioconductor.org).^dFour groups of differentially expressed genes in relation to dormant spores were reported.^eGenes were grouped into three clusters according to their expression profiles by K-means algorithm.^fNumber of gene orthologs in each of the hierarchical clusters.

(*P* = 0.030). These orthologs over-represented in Max1 cluster belonged to diverse FunCat categories (Supplementary Data S3). Both the *D.discoideum* cluster III and the *N.crassa* cluster Max1 comprised genes that showed an expression nadir in dormant spores, but maximum expression during spore swelling. Thus, a cross-species examination of gene expression patterns during spore germination in *N.crassa*, *U.maydis* and *D.discoideum* indicated that an evolutionarily conserved mechanism may underlie the breaking of dormancy and the initiation of transcription of certain sets of genes involved in early events of the germination process. Continued comparative analysis of transcriptional profiles of the germination processes in diverse organisms may further extend these observations.

CONCLUSION

This study describes the transcriptional program underlying the biochemical and physiological events associated with conidial germination in the filamentous fungus, *N.crassa*. To observe this transcriptional regulation on a genomic scale, we developed a 70mer oligonucleotide microarray for transcriptional profiling and the evaluated its performance based on the extensive biochemical literature on the conidial germination process (13,19,74). Our gene set for oligonucleotide design comprises 10 032 genes that were predicted using a combination of FGENESH, FGENESH+ and Gene-wise by the Broad Institute (6) and by annotation via FGENESH, plus manual corrections via GENEMARK, GENSCAN and GENEFINDER by MIPs (7). Out of 3366 oligonucleotides that we designed, synthesized and robotically deposited for microarray analysis, we obtained hybridization data for 3054 for at least one time point between conidial

germination until 16 h of growth, indicating that gene predictions of the *Neurospora* genome are of high quality. We obtained strong statistical estimates of expression levels for 1287 genes during the process of conidial germination. Our estimates of gene expression levels from transcriptional profiling of conidial germination are remarkably consistent with previous data assessing transcript levels of a number of genes and with biochemical processes that have been associated with germination, even though different laboratory strains and conditions were used. These observations reject the hypothesis of non-concordance between mRNA abundance and protein or cellular activities.

N.crassa has an advantage over numerous other eukaryotic microorganisms because biochemical mechanisms associated with the germination process have been extensively studied [for review, see (13,19,74)]. Cross-species examination of expression profiles during spore germination in *N.crassa*, *U.maydis* and *D.discoideum* revealed that some of the same sets of orthologous genes were activated or deactivated under comparable developmental stages among these diverse organisms (40,41). These observations suggest that evolutionary conserved mechanisms are associated with spore germination in general. Orthologous genes with similar profiles during germination belonged to diverse functional categories. This may indicate that common regulatory features that affect diverse processes are conserved during spore germination across diverse species. A comparison between additional complete and robust transcriptional datasets on spore germination in diverse species will further test this hypothesis and may illuminate common regulatory processes.

Our power to determine the expression levels of genes relevant to conidial germination was maximized by a highly replicated, closed circuit experimental design. Closed loops of experimental comparisons within an experimental design

have been advocated for a variety of reasons (29,30,42), but primarily because they allow the statistical integration of information from both transitive and direct comparisons (46,47) during estimation of gene expression levels from transcriptional profiling data. Expression levels at time points during conidial germination were estimated by a Bayesian analysis (30) that integrates all the ratio data from a replicated circuit design to yield gene expression level estimates, 95% credible intervals and Bayesian posterior probabilities for directional differential expression. The method is robust to small amounts of missing data, and with sufficient replication allows detection of both large and small differences in gene expression (45).

The *N.crassa* genome project revealed that 41% of predicted protein-coding sequences have no similarity to known sequences (6). Functional characterization of such genes is one of the most important tasks for the post genome area of *N.crassa*. Of the 1287 genes for which strong estimates of gene expression level were acquired, 473 have been described as hypothetical, conserved hypothetical or putative genes. Transcriptional profiles from these 473 genes were distributed among clusters associated with initiation (0.5–1 h), early (1–4 h), middle (4–8 h) and late (12–16 h) germination processes. The cascade of biochemical processes associated with conidial germination has been described in detail (13); transcriptional profiles of some of these unknown genes correlate well with biochemical processes associated with conidial germination. Thus, microarray data in *N.crassa* will guide future laboratory experiments with regard to functional annotation of hypothetical genes. The availability of full genome microarrays will allow expression patterns associated with developmental states in a filamentous fungus to be discerned and will aid in the functional analysis of deletion mutants generated by the *Neurospora* functional genomics project (<http://www.dartmouth.edu/~neurosporagenome/>).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Sarah Calvo and James Galagan (The Broad Institute) for making genomic data available for oligonucleotide prediction, Qikai Xu and Gad Shaulsky for sharing *Dictyostelium* microarray data with us. We thank Audrey Gasch, Mike Eisen, Jing Zhu and Joe DeRisi for valuable discussions on the development of oligonucleotide microarrays for *N.crassa*. We thank Sarah Brown, Stu Brody and Matt Sachs for helpful suggestions and for reviewing our manuscript. The work was funded by a University of California Biological Sciences award to N.L.G, a National Institutes of Health multi-institutional program project grant (GM068087) to N.L.G (Core III: transcriptional profiling) and a National Science Foundation grant (DEB-0316710) to J.W.T. J.P.T. was supported by a Miller Post-doctoral Fellowship. Funding to pay the Open Access publication charges for this article was provided by NIH grant GM068087 to N.L.G.

Conflict of interest statement. None declared.

REFERENCES

- Collins,F.S., Lander,E.S., Rogers,J., Waterston,R.H. and Conso,I.H.G.S. (2004) Finishing the euchromatic sequence of the human genome. *Nature*, **431**, 931–945.
- Davis,R.H. (2000) *NEUROSPORA: Contributions of a Model Organism*. Oxford University Press, NY.
- Davis,R.H. and Perkins,D.D. (2002) *Neurospora*: a model of model microbes. *Nature Rev. Genet.*, **3**, 397–403.
- Borkovich,K.A., Alex,L.A., Yarden,O., Freitag,M., Turner,G.E., Read,N.D., Seiler,S., Bell-Pedersen,D., Paietta,J., Plesofsky,N. *et al.* (2004) Lessons from the genome sequence of *Neurospora crassa*: tracing the path from genomic blueprint to multicellular organism. *Microbiol. Molec. Biol. Rev.*, **68**, 1–108.
- Bistis,G.N., Perkins,D.D. and Read,N.D. (2003) Cell types of *Neurospora crassa*. *Fungal Genet. Newslett.*, **50**, 17–19.
- Galagan,J.E., Calvo,S.E., Borkovich,K.A., Selker,E.U., Read,N.D., Jaffe,D., FitzHugh,W., Ma,L.J., Smirnov,S., Purcell,S. *et al.* (2003) The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature*, **422**, 859–868.
- Mannhaupt,G., Montrone,C., Haase,D., Mewes,H.W., Aign,V., Hoheisel,J.D., Fartmann,B., Nyakatura,G., Kempken,F., Maier,J. *et al.* (2003) What's in the genome of a filamentous fungus? Analysis of the *Neurospora* genome sequence. *Nucleic Acids Res.*, **31**, 1944–1954.
- DeRisi,J., Penland,L., Brown,P.O., Bittner,M.L., Meltzer,P.S., Ray,M., Chen,Y., Su,Y.A. and Trent,J.M. (1996) Use of a cDNA microarray to analyze gene expression patterns in human cancer. *Nature Genet.*, **14**, 457–460.
- Eisen,M. and Brown,P.O. (1999) DNA arrays for analysis of gene expression. In Weissman,S.M. (ed.), *Methods in Enzymology*. Academic Press, San Diego, CA. Vol. 303, pp. 179–205.
- Wu,L.F., Hughes,T.R., Davierwala,A.P., Robinson,M.D., Stoughton,R. and Altschuler,S.J. (2002) Large-scale prediction of *Saccharomyces cerevisiae* gene function using overlapping transcriptional clusters. *Nature Genet.*, **31**, 255–265.
- Pradet-Balade,B., Boulme,F., Beug,H., Mullner,E.W. and Garcia-Sanz,J.A. (2001) Translation control: bridging the gap between genomics and proteomics?. *Trends Biochem. Sci.*, **26**, 225–229.
- Feder,M.E. and Walser,J.C. (2005) The biological limitations of transcriptomics in elucidating stress and stress responses. *J. Evol. Biol.*, **18**, 901–910.
- Schmit,J.C. and Brody,S. (1976) Biochemical Genetics of *Neurospora crassa* Conidial Germination. *Bacteriol. Rev.*, **40**, 1–41.
- Roca,M.G., Arlt,J., Jeffree,C.E. and Read,N.D. (2005) Cell biology of conidial anastomosis tubes in *Neurospora crassa*. *Eukaryotic Cell*, **4**, 911–919.
- Springer,M.L. (1993) Genetic control of fungal differentiation—the 3 sporulation pathways of *Neurospora crassa*. *Bioessays*, **15**, 365–374.
- Lowry,R.J., Durkee,T.L. and Sussman,A. (1967) Ultrastructural studies of microconidial formation in *Neurospora crassa*. *J. Bacteriol.*, **94**, 1757–1763.
- Denning,D.W., Anderson,M.J., Turner,G., Latge,J.P. and Bennett,J.W. (2002) Sequencing the *Aspergillus fumigatus* genome. *Lancet Infect. Dis.*, **2**, 251–253.
- Griffin,D.H. (1994) *Fungal Physiology*, 2nd edn. Wiley-Liss, Inc., NY.
- d'Enfert,C. (1997) Fungal spore germination: Insights from the molecular genetics of *Aspergillus nidulans* and *Neurospora crassa*. *Fungal Genet. Biol.*, **21**, 163–172.
- d'Enfert,C., Bonini,B.M., Zapella,P.D.A., Fontaine,T., da Silva,A.M. and Terenzi,H.F. (1999) Neutral trehalases catalyze intracellular trehalose breakdown in the filamentous fungi *Aspergillus nidulans* and *Neurospora crassa*. *Molec. Microbiol.*, **32**, 471–483.
- Mirkes,P.E. and McCalley,B. (1976) Synthesis of polyadenylic acid-containing ribonucleic acid during the germination of *Neurospora crassa* conidia. *J. Bacteriol.*, **125**, 174–180.
- Loo,M. (1975) *Neurospora crassa* temperature-sensitive mutant apparently defective in protein synthesis. *J. Bacteriol.*, **121**, 286–295.
- Momany,M. (2002) Polarity in filamentous fungi: establishment, maintenance and new axes. *Curr. Opin. Microbiol.*, **5**, 580–585.
- Pandey,A., Roca,M.G., Read,N.D. and Glass,N.L. (2004) Role of a mitogen-activated protein kinase pathway during conidial germination and hyphal fusion in *Neurospora crassa*. *Eukaryotic Cell*, **3**, 348–358.

25. Vogel, H.J. (1956) A convenient growth medium for *Neurospora* (Medium N). *Microbial. Genet. Bull.*, **13**, 42–43.
26. Wang, H.-Y., Malek, R.L., Kwitek, A.E., Greene, A.S., Luu, T.V., Behbahani, B., Frank, B., Quackenbush, J. and Lee, N.H. (2003) Assessing unmodified 70-mer oligonucleotide probe performance on glass-slide microarrays. *Genome Biol.*, **4**, R5.1–R5.13.
27. Bozdech, Z., Zhu, J., Joachimiak, M.P., Cohen, F.E., Pulliam, B. and DeRisi, J.L. (2003) Expression profiling of the schizont and trophozoite stages of *Plasmodium falciparum* with a long-oligonucleotide microarray. *Genome Biol.*, **4**, R9.
28. Martinez, M.J., Aragon, A.D., Rodriguez, A.L., Weber, J.M., Timlin, J.A., Sinclair, M.B., Haaland, D.M. and Werner-Washburne, M. (2003) Identification and removal of contaminating fluorescence from commercial and in-house printed DNA microarrays—art. no. e18. *Nucleic Acids Res.*, **31**, E18.
29. Yang, Y.H. and Speed, T. (2002) Design issues for cDNA microarray experiments. *Nature Rev. Genet.*, **3**, 579–588.
30. Townsend, J.P. and Hartl, D.L. (2002) Bayesian analysis of gene expression levels: statistical quantification of relative mRNA level across multiple strains or treatments. *Genome Biol.*, **3**, research0071.0071–0071.0016.
31. Eisen, M.B., Spellman, P.T., Brown, P.O. and Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns. *Proc. Natl Acad. Sci. USA*, **95**, 14863–14868.
32. Seo, J. and Shneiderman, B. (2002) Interactively exploring hierarchical clustering results. *IEEE Comput.*, **35**, 80–86.
33. Ruepp, A., Zollner, A., Maier, D., Albermann, K., Hani, J., Mokrejs, M., Tetko, I., Guldener, U., Mannhaupt, G., Munsterkotter, M. *et al.* (2004) The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. *Nucleic Acids Res.*, **32**, 5539–5545.
34. Frishman, D., Albermann, K., Hani, J., Heumann, K., Metanowski, A., Zollner, A. and Mewes, H.W. (2001) Functional and structural genomics using PEDANT. *Bioinformatics*, **17**, 44–57.
35. McGuire, A.M., Hughes, J.D. and Church, G.M. (2000) Conservation of DNA regulatory motifs and discovery of new motifs in microbial genomes. *Genome Res.*, **10**, 744–757.
36. Liu, X.S., Brutlag, D.L. and Liu, J.S. (2002) An algorithm for finding protein-DNA binding sites with applications to chromatin-immunoprecipitation microarray experiments. *Nat. Biotechnol.*, **20**, 835–839.
37. Bailey, T.L. and Elkan, C. (1994) *Second International Conference on Intelligent Systems for Molecular Biology*. AAAI Press, Menlo Park, CA.
38. Kreppel, L., Fey, P., Gaudet, P., Just, E., Kibbe, W.A., Chisholm, R.L. and Kimmel, A.R. (2004) dictyBase: a new *Dictyostelium discoideum* genome database. *Nucleic Acids Res.*, **32**, D332–D333.
39. Eichinger, L., Pachebat, J.A., Glockner, G., Rajandream, M.A., Sugang, R., Berriman, M., Song, J., Olsen, R., Szafranski, K., Xu, Q. *et al.* (2005) The genome of the social amoeba *Dictyostelium discoideum*. *Nature*, **435**, 43–57.
40. Zahiri, A.R., Babu, M.R. and Saville, B.J. (2005) Differential gene expression during teliospore germination in *Ustilago maydis*. *Mol. Genet. Genomics*, **273**, 394–403.
41. Xu, Q.-K., Ibarra, M., Mahadeo, D., Shaw, C., Huang, E.Y., Kuspa, A., Cotter, D. and Shaulsky, G. (2004) Transcriptional transitions during *Dictyostelium* spore germination. *Eukaryotic Cell*, **3**, 1101–1110.
42. Kerr, M.K. and Churchill, G.A. (2001) Statistical design and the analysis of gene expression microarray data. *Genet. Res.*, **77**, 123–128.
43. Wolfinger, R.D., Gibson, G., Wolfinger, E.D., Bennett, L., Hamadeh, H., Bushel, P., Afshari, C. and Paules, R.S. (2001) Assessing gene significance from cDNA microarray expression data via mixed models. *J. Comput. Biol.*, **8**, 625–637.
44. Jin, W., Riley, R.M., Wolfinger, R.D., White, K.P., Passador-Gurgel, G. and Gibson, G. (2001) The contributions of sex, genotype and age to transcriptional variance in *Drosophila melanogaster*. *Nature Genet.*, **29**, 389–395.
45. Townsend, J.P. (2004) Resolution of large and small differences in gene expression using models for the Bayesian analysis of gene expression levels and spotted DNA microarrays. *BMC Bioinformatics*, **5**, 54.
46. Townsend, J.P. (2003) Multifactorial experimental design and the transitivity of ratios with spotted DNA microarrays. *BMC Genomics*, **4**, 41.
47. Vinciotti, V., Khanin, R., D'Alimonte, D., Liu, X., Cattini, N., Hotchkiss, G., Bucca, G., de Jesus, O., Rasaiyaah, J., Smith, C.P. *et al.* (2005) An experimental evaluation of a loop versus a reference design for two-channel microarrays. *Bioinformatics*, **21**, 492–501.
48. Townsend, J.P. (2005) In Zimmer, E. and Roalson, E. (eds), *Methods in Enzymology: Molecular Evolution, Producing the Biochemical Data, Part B*. Academic Press, San Diego, Vol. 395, pp. 597–617.
49. Sachs, M.S. and Yanofsky, C. (1991) Developmental expression of genes involved in conidiation and amino acid biosynthesis in *Neurospora crassa*. *Dev. Biol.*, **148**, 117–128.
50. Nelson, M.A., Kang, S.C., Braun, E.L., Crawford, M.E., Dolan, P.L., Leonard, P.M., Mitchell, J., Armijo, A.M., Bean, L., Blueyes, E. *et al.* (1997) Expressed sequences from conidial, mycelial, and sexual stages of *Neurospora crassa*. *Fungal Genet. Biol.*, **21**, 348–363.
51. Dolan, P.L., Natvig, O.D. and Nelson, M.A. (2000) *Neurospora* Proteome 2000. *Fungal Genet. Newsl.*, **47**, 7–24.
52. DeRisi, J., Iyer, V.R. and Brown, P.O. (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science*, **278**, 680–686.
53. Berlin, V. and Yanofsky, C. (1985) Isolation and characterization of genes differentially expressed during conidiation of *Neurospora crassa*. *Mol. Cell. Biol.*, **5**, 849–855.
54. Bieszke, J.A., Braun, E.L., Bean, L.E., Kang, S.C., Natvig, D.O. and Borkovich, K.A. (1999) The nop-1 gene of *Neurospora crassa* encodes a seven transmembrane helix retinal-binding protein homologous to archaeal rhodopsins. *Proc. Natl Acad. Sci. USA*, **96**, 8034–8039.
55. Fracella, F., Scholle, C., Kallies, A., Haker, T., Schroder, T. and Rensing, L. (1997) Differential HSC70 expression during asexual development of *Neurospora crassa*. *Microbiology*, **143**, 3615–3624.
56. Schmit, J.C. and Brody, S. (1975) *Neurospora crassa* conidial germination: the role of endogenous amino acid pools. *J. Bacteriol.*, **124**, 232–242.
57. Scarborough, G.A. (1973) Transcript in *Neurospora*. *Int. Rev. Cytol.*, **34**, 103–122.
58. Colvin, H.J., Sauer, B.L. and Munkres, K.D. (1973) Respiration of wild type and extrachromosomal mutants of *Neurospora crassa*. *J. Bacteriol.*, **116**, 1314–1321.
59. Xie, X., Wilkinson, H.H., Correa, A., Lewis, Z.A., Bell-Pedersen, D. and Ebbole, D.J. (2004) Transcriptional response to glucose starvation and functional analysis of a glucose transporter of *Neurospora crassa*. *Fungal Genet. Biol.*, **41**, 1104–1119.
60. Schmit, J.C. and Brody, S. (1975) Changes in glucosamine and galactosamine levels during conidial germination in *Neurospora crassa*. *J. Bacteriol.*, **122**, 1062–1070.
61. Ruiz-Herrera, J. (1992) *Fungal Cell Wall: Structure, Synthesis and Assembly*. Boca Raton, Florida.
62. Gasch, A.P., Moses, A.M., Chiang, D.Y., Fraser, H.B., Berardini, M. and Eisen, M.B. (2004) Conservation and evolution of cis-regulatory systems in ascomycete fungi. *Plos Biol.*, **2**, 2202–2219.
63. Sinha, S. and Tompa, M. (2002) Discovery of novel transcription factor binding sites by statistical overrepresentation. *Nucleic Acids Res.*, **30**, 5549–5560.
64. Birnbaum, K., Benfey, P.N. and Shasha, D.E. (2001) cis element/transcription factor analysis (cis/TF): a method for discovering transcription factor/cis element relationships. *Genome Res.*, **11**, 1567–1573.
65. Liu, X., Brutlag, D.L. and Liu, J.S. (2001) BioProspector: discovering conserved DNA motifs in upstream regulatory regions of co-expressed genes. *Pac. Symp. Biocomput.*, 127–138.
66. Shi, Y.G. and Tyler, B.M. (1991) Coordinate expression of ribosomal-protein genes in *Neurospora crassa* and identification of conserved upstream sequences. *Nucleic Acids Res.*, **19**, 6511–6517.
67. Cujec, T.P. and Tyler, B.M. (1996) Functional promoter elements common to ribosomal protein and ribosomal RNA genes in *Neurospora crassa*. *Mol. Gen. Genet.*, **253**, 205–216.
68. Cujec, T.P. and Tyler, B.M. (1996) Nutritional and growth control of ribosomal protein mRNA and rRNA in *Neurospora crassa*. *Nucleic Acids Res.*, **24**, 943–950.
69. Yan, W., Schilke, B., Pfund, C., Walter, W., Kim, S.W. and Craig, E.A. (1998) Zuo1, a ribosome-associated DnaJ molecular chaperone. *EMBO J.*, **17**, 4809–4817.

70. Pfund, C., Lopez-Hoyo, N., Ziegelhoffer, T., Schilke, B.A., Lopez-Buesa, P., Walter, W.A., Wiedmann, M. and Craig, E.A. (1998) The molecular chaperone Ssb from *Saccharomyces cerevisiae* is a component of the ribosome nascent chain complex. *EMBO J.*, **17**, 3981–3989.
71. Hundley, H.A., Walter, W., Bairstow, S. and Craig, E.A. (2005) Human Mpp11 J protein: ribosome-tethered molecular chaperones are ubiquitous. *Science*, **308**, 1032–1034.
72. Salmi, M.L., Bushart, T.J., Stout, S.C. and Roux, S.J. (2005) Profile and analysis of gene expression changes during early development in germinating spores of *Ceratopteris richardii*. *Plant Physiol.*, **138**, 1734–1745.
73. Soeda, Y., Konings, M.C.J.M., Vorst, O., van Houwelingen, A.M.M.L., Stoop, G.M., Maliapaard, C.A., Kodde, J., Bino, R.J., Groot, S.P.C. and van der Geest, A.H.M. (2005) Gene expression programs during *Brassica oleracea* seed maturation, osmopriming, and germination are indicators of progression of the germination process and the stress tolerance level. *Plant Physiol.*, **137**, 354–368.
74. Oshero, N. and May, G.S. (2001) The molecular mechanisms of conidial germination. *Fems Microbiol. Lett.*, **199**, 153–160.